

SPECIFICATION

Title of the Invention :

人物の顔の検出方法およびその装置

Inventor :

Shin YAMADA
Ryuta ITOH
Kenji NAGAO
Masaki SOUMA

人物の顔の検出方法およびその装置

発明の背景

発明分野

本発明は、人間の顔や人物に関する情報をキーにして動画像を検索する動画像
5 検索装置や、自動車の運転手や乗員を監視する監視システムや、カメラに映った
顔をデータベース上の顔と認識する顔認識システムに係り、画像中から人物の顔
の領域、顔の情報などを検出する技術に関するものである。

先行技術

従来から、人物の顔を検出する技術が開発されており、例えば特開平 7-31
10 1833 号公報に開示されている顔検出装置がある。以下、従来の顔検出装置に
ついて図 22 を用いて説明する。

従来の顔検出装置 2220 は、目と口に注目したものであり、輝度が局所的に
最も小さくなる輝度極小点と輝度が大きくなっていく輝度変曲点とを検出して、
この 2 点に挟まれる領域を顔の構成要素の領域として切り出す領域検出装置 22
15 21 と、顔構成要素領域の大きさや位置関係より顔の候補を検出する顔候補検出
装置 2222 と、顔候補を詳しく調べて顔であるか否かを判定する顔判定装置 2
223 との 3 つの処理装置から構成される。入力画像信号 2231 は、人物の顔
を含む画像である。

入力画像信号 2231 を一次微分手段 2201 により、画像上の上から下の方
20 向に一次微分を行い、一次微分信号 2232 を出力する。一次微分信号 2232
を二値化手段 2202 により 0 で二値化し、一次微分二値化信号 2233 を出力
する。また入力画像信号 2231 を二次微分手段 2203 により二次微分を行い
、二次微分信号 2234 を出力する。二次微分信号 2234 を二値化手段 220
4 により 0 で二値化し、二次微分二値化信号 2235 を出力する。

25 論理和手段 2205 により、一次微分二値化信号 2233 と二次微分二値化信
号 2235 の論理和をとり、目・口第一候補信号 2236 を出力する。連結領域
特徴量算出手段 2206 は、目・口第一候補信号 2236 に対して、連結領域を
成している各領域に対してその面積、重心位置、領域の縦・横の長さ、また入力
画像信号 2231 をも入力とし、各領域に対して領域内の輝度の平均・分散の領

域の特徴量を検出し、領域特徴量信号 2 2 3 7 として出力する。

領域特徴量信号が入力した目第二候補判定手段 2 2 0 7 が、各領域の面積、領域の縦・横の長さ、輝度平均・分散を調べて、前記領域が目らしい領域を判定し、領域の特徴量と合わせて目第二候補信号 2 2 3 8 として出力する。同様に前記

- 5 領域特徴量信号を入力とする口第二候補判定手段 2 2 0 8 が、各領域の面積、領域の縦・横の長さ、輝度平均・分散を調べて、前記領域が口らしい領域を判定し、領域の特徴量と合わせて口第二候補信号 2 2 3 9 として出力する。

- 顔候補判定手段 2 2 0 9 により、前記目第二候補信号から目候補領域を 2 つ、前記口第二候補信号から口候補領域を 1 つ、すべての領域が互いに重複しないように選択し、各領域の重心位置を調べて顔らしい配置をしている候補の組をすべての組み合わせについて調べ、顔候補信号 2 2 4 0 として出力する。

- 顔候補領域画像切り出し手段 2 2 1 0 は、アフィン変換を用いて、該顔候補信号中の左右の目の候補領域の重心位置を基準に、顔の存在する候補領域を切り出し、顔候補画像信号 2 2 4 1 として出力する。顔判定手段 2 2 1 1 は、前記顔候補画像信号と顔標準パターンとの間の距離を算出し、ある閾値よりも小さければ、前記入力信号の対応する場所に人間の顔が撮影されていると判断し、顔の存在する位置・大きさ・角度を顔信号 2 2 4 2 として出力する。

- 以上のように、従来の技術では、輝度が局所的に最も小さくなる輝度極小点と輝度が大きくなっていく輝度変曲点とを検出して、この 2 点に挟まれる領域を目や口の候補領域として切り出してから、これらの候補領域の形状特徴と輝度特徴とから目の候補及び口の候補を検出し、目の候補と口の候補との位置関係より顔の候補を検出し、顔候補を詳しく調べて顔であるか否かを判定する。

- しかし、顔以外の背景物が多く映っている映像の場合には、輝度極小点と輝度変曲点とが多数存在するために、目や口の候補領域が多数検出され、多くの誤検出が発生するという課題を有していた。

従来技術では、目や口の候補領域の形状特徴と輝度特徴とから目の候補と口の候補とを検出している。しかし、個人差や表情の変化によって目と口の形状が大きく変化するため、目や口の候補領域として誤検出された背景部分は、目または口の候補領域として多数残ってしまう。さらに、斜め向きの顔や傾いた顔も検出

するように設計した場合に顕著であるが、目の候補と口の候補との位置関係を使って顔の候補を検出するときに、誤検出された背景部分が顔の候補として多数残る。顔候補を詳しく調べる処理で誤検出を何%抑制できるかは、顔候補を詳しく調べる処理のアルゴリズムと閾値に依存して変化し、顔の候補として誤検出される背景部分は少ない方が望ましいだけでなく、多くのアルゴリズムでは計算量が急激に増大してしまう。

また、従来方式では、ヒゲが口にかかった顔を検出しようとしても、ヒゲと口を分離できずに検出もれになる場合があるという課題を有していた。

発明の要旨

本発明は、目や口の動きを含む顔の表情の変化や、ヒゲの有無による顔の変動があっても少ない計算量で誤検出を抑制でき、人物の顔の領域などの情報を検出するのに好適な検出装置及び方法を提供することを目的とする。

本発明は、画像中から目の候補領域と目の間の候補領域とを検出し、目の候補領域と目の間の候補領域との位置関係から目の間の領域を決定し、目の間の領域を含むように顔の領域を決定するものである。

これにより、目の間の領域は顔の表情による変動が少なく、口やヒゲと離れているので、目や口の動きを含む顔の表情やヒゲの有無による顔の変動に強く、人物の顔の領域などの情報を検出することができる。

また、複数の画素が連結領域かどうかを判定するのではなく、最初に小領域に分割してから小領域が目や目の間の領域であるかどうかを判定することにより、少ない計算量で人物の顔の領域などの情報を検出できる。

さらに、目の間の領域の重なりを調べたり、目の間の候補領域内の特徴や目の候補領域内の特徴を判定して目の間の領域を決定することで、顔以外の背景物が多く映っている場合を含めて、少ない計算量で人物の顔の領域を高精度で検出することができる。

図面の簡単な説明

図1は、本発明の実施の形態1における人物の顔の検出装置の構成図

図2は、本発明の実施の形態1における人物の顔の検出装置の動作を示すフローチャート

図3(a)は小領域の例を示す図

図3(b)は検出した人物の顔の候補領域と目の間の候補領域の例を示す図

5 図3(c)は検出した人物の顔の領域と目の間の領域の例を示す図

図4は、本発明の実施の形態2における人物の顔の検出装置の構成図

図5は、目の間の領域の検出対象となる2つの領域の例を示す図

図6は、本発明の実施の形態2における人物の顔の検出装置の動作を示すフローチャート

10 図7は、本発明の実施の形態2における2つの検出された目の間の領域の例を示す図

図8は、本発明の実施の形態3における人物の顔の検出装置の構成図

図9は、本発明の実施の形態3における人物の顔の検出装置の動作を示すフローチャート

15 図10は、目の間の候補領域と左目の候補領域と右目の候補領域の例を示す図

図11は、本発明の実施の形態4における人物の顔の検出装置の構成図

図12は、本発明の実施の形態4における目決定手段の構成図

図13は、本発明の実施の形態4における人物の顔の検出装置の動作を示すフローチャート

20 図14は、目の間の候補領域と左目の候補領域と右目の候補領域の例を示す図

図15は、本発明の実施の形態5における人物の顔の検出装置の構成図

図16は、本発明の実施の形態5における人物の顔の検出装置の動作を示すフローチャート

25 図17は、本発明の実施の形態6における動画像検索装置の構成図

図18は、本発明の実施の形態6における動画像検索装置の動画像情報ファイルと顔画像ファイルを作成する動作を示すフローチャート

図19は、本発明の実施の形態6における動画像検索装置の顔画像を検索する動作を示すフローチャート

図20は、本発明の実施の形態7における顔認識装置の構成図

図21は、本発明の実施の形態7における顔認識装置の動作を示すフローチャート

図22は、顔検出装置を示す構成図

5

望ましい実施態様

以下に、本発明の実施の形態について、図を用いて説明する。

(実施の形態1)

図1は、本発明の実施の形態1における人物の顔の検出装置の構成図である。

図1において、人物の顔の検出装置122には入力画像信号102が入力され、
10 入力画像信号102がそのまま目間領域検出手段101の分割手段103に入力される。

分割手段103は、入力された入力画像信号102の画像全体を複数個の小領域に分割し、各小領域の画像を小領域画像信号104として順次出力する。

エッジ特徴量計算手段105は、小領域画像信号104を読み込んで、小領域
15 における各画素のエッジ強度の総和をエッジ特徴量信号106として出力する。

輝度ヒストグラム計算手段107は、小領域画像信号104を読み込んで、小領域内の画素のヒストグラムを輝度ヒストグラム信号108として出力する。

輝度ヒストグラム比較手段109は、輝度ヒストグラム信号108を読み込んで、水平方向に隣接する2つの小領域間の輝度ヒストグラムの差を輝度ヒストグラム差信号110として出力する。
20

第1のバッファ手段111は、複数の小領域のエッジ特徴量信号106と輝度ヒストグラム差信号110を記憶する。

目候補検出手段112は、第1のバッファ手段111に記憶したエッジ特徴量信号と輝度ヒストグラム差信号を読み出して、目の候補となる小領域（目候補）
25 を検出し、目の候補となる小領域の座標を目候補信号113として出力する。

目間候補検出手段114は、第1のバッファ手段111からエッジ特徴量信号106と輝度ヒストグラム差信号110を読み出して、目の間の候補となる小領域（目間候補）を検出し、目の間の候補となる小領域の座標を表す目間候補信号115と眼鏡の有無を表す眼鏡判定信号116を出力する。

第2のバッファ手段117は、目候補信号113と目間候補信号115と眼鏡判定信号116を記憶する。

目間決定手段118は、第2のバッファ手段117から目候補信号113と目間候補信号115と眼鏡判定信号116を読み出し、必要に応じて第1のバッファ手段111から輝度ヒストグラム差信号110を読み出し、2つの目の候補の小領域の間に目の間の候補の小領域が存在するかどうかを調べて目の間の領域を決定し、目の間の領域の座標を目間領域信号119として出力する。

顔決定手段120は、目間領域信号119を読み込んで、顔の領域の座標を表す顔領域信号121を出力する。

10 なお、図1の人物の顔の検出装置122は、パソコンのCPUとメモリを用いて人物の顔の検出プログラムによって実現することもできる。

以下、図1の人物の顔の検出装置122の動作について、図2のフローチャートを用いて説明する。

手順201では、分割手段103が、画像入力信号102によって入力された
15 画像全体をVslit×Hslitのサイズ（Vslit、Hslitは1以上の整数）の小領域に分割し、各小領域の画像を小領域画像信号104として順次出力する。小領域は、図3（a）に示すように、例えばVslit=12、Hslit=2（Vslit、Hslitはあらかじめ設定された値）とする。

手順202では、エッジ特徴量計算手段105が、小領域画像信号104を読み込んで、小領域内の画素(x,y)に対して（x、yは1以上の整数）垂直方向エッジ強度の二乗 $Ve(x,y)$ を

$$Ve(x,y)=[Y(x,y-1)-Y(x,y+1)] \times [Y(x,y-1)-Y(x,y+1)]$$

によって求め、左からi番目、上からj番目の小領域(i,j)内部の $Ve(x,y)$ の総和をSve(i,j)とする（i、jは1以上の整数）。ただし、 $Y(x,y)$ は画素(x,y)の輝度
25 を表し、0以上255以下の整数である。左からi番目、上からj番目の小領域(i,j)における各画素のエッジ強度の総和Sve(i,j)をエッジ特徴量信号106として出力する。

手順203では、輝度ヒストグラム計算手段107が、小領域画像信号104を読み込んで、小領域(i,j)内部の画素(x,y)に対して輝度値の頻度を調べてヒス

トグラム $H(i,j,bin)$ を作成する。ただし、本実施例では、上位3ビットの値の頻度を調べてヒストグラム $H(i,j,bin)$ を作成するものとする (bin は1以上8以下の整数)。左から i 番目、上から j 番目の小領域 (i,j) 内の画素のヒストグラム $H(i,j,bin)$ を輝度ヒストグラム信号108として出力する。

- 5 手順204では、輝度ヒストグラム比較手段109が、輝度ヒストグラム信号108を読み込んで、水平方向に隣接する2つの小領域 (i,j) 、 $(i-1,j)$ の間の輝度ヒストグラムの差 $Dh(i,j)$ を、

$$Dh(i,j) = \sum [[H(i,j,bin) - H(i-1,j,bin)] \times [H(i,j,bin) - H(i-1,j,bin)]]$$

- 10 によって求める。左から i 番目、上から j 番目の小領域 (i,j) と、左から $(i-1)$ 番目、上から j 番目の小領域 $(i-1,j)$ との間の輝度ヒストグラムの差 $Dh(i,j)$ を、輝度ヒストグラム差信号110として出力する。

手順205では、第1のバッファ手段111が、複数の小領域のエッジ特徴量信号106と輝度ヒストグラム差信号110を記憶する。

- 15 手順206では、目の部分で水平に近い方向の線分が多いので、エッジ特徴量信号 $Sve(i,j)$ と第1の閾値 $Th1$ を比較し、

$$Sve(i,j) > Th1$$

を満足する小領域があらかじめ設定した個数 $N1$ 個以上水平方向に連続し、さらに垂直に近い方向の線分が少ないので、その中であらかじめ設定した個数 $N2$ 個

- 20 以上の小領域が、輝度ヒストグラム差信号 $Dh(i,j)$ と第2の閾値 $Th2$ を比較する式

$$Dh(i,j) < Th2$$

を満足するとき、この $N1$ 個の小領域を目候補とする。そして、目候補の左端の小領域の左上の座標と、目候補の右端の小領域の右下の座標を、目候補信号113として出力する。

- 25 手順207では、眼鏡をかけていない顔の目の間で、水平に近い方向の線分が少ないので、エッジ特徴量信号 $Sve(i,j)$ と第3の閾値 $Th3$ を比較し、

$$Sve(i,j) < Th3$$

を満足する小領域があらかじめ設定した個数 $N3$ 個以上水平方向に連続し、さらに垂直に近い方向の線分も少ないので、その中であらかじめ設定した個数 $N4$ 個

以上の小領域が、輝度ヒストグラム差信号 $Dh(i, j)$ と第4の閾値 $Th4$ を比較する式

$$Dh(i, j) < Th4$$

を満足するとき、このN3個の小領域を目間候補とする。そして、目間候補の左端の小領域の左上の座標と、目間候補の右端の小領域の右下の座標を、目間候補

- 5 信号115として出力する。また、この目間候補信号に対応する眼鏡判定信号116を「眼鏡なし」にして出力する。

手順208では、眼鏡をかけている顔の目間に水平に近い方向の線分が多いので、エッジ特徴量信号 $Sve(i, j)$ と第5の閾値 $Th5$ の比較と、エッジ特徴量信号 $Sve(i, j-1)$ と第6の閾値 $Th6$ の比較を実行し、

10 $Sve(i, j) > Th5$

$$Sve(i, j-1) < Th6$$

を満足する小領域の組 $(i, j), (i, j-1)$ があらかじめ設定した個数N5個以上水平方向に連続し、さらに額部分に線分が少ないので、その中であらかじめ設定した個数N6個以上の小領域が、輝度ヒストグラム差信号 $Dh(i, j-1)$ と第7の閾値 $Th7$

- 15 を比較する式

$$Dh(i, j-1) < Th7$$

を満足するとき、N5個の小領域の組の下半分を目間候補とする。そして、目間候補の左端の小領域の左上の座標と、目間候補の右端の小領域の右下の座標を、目間候補信号115として出力する。また、この目間候補信号に対応する眼鏡判

- 20 定信号116を「眼鏡あり」にして出力する。

手順209では、第2のバッファ手段117が、目候補信号113と目間候補信号115と眼鏡判定信号116を記憶する。

手順210では、目間決定手段118が、第2のバッファ手段117から目候補信号113と目間候補信号115を読み出し、水平に並ぶ2つの目候補に挟ま

- 25 れる目間候補を検出する。

検出された目間候補に対応する眼鏡判定信号が「眼鏡なし」である場合には、検出された目間候補の左側にある目候補の目候補信号の右下の座標と、検出された目間候補の右側にある目候補の目候補信号の左上の座標を、目間領域信号118として出力する。検出された目間候補に対応する眼鏡判定信号が「眼鏡あり」

である場合には、検出された目間候補の左側にある目候補の目候補信号の右下の座標と、検出された目間候補の右側にある目候補の目候補信号の左上の座標に囲まれる領域を目間領域の候補とし、その中の小領域の輝度ヒストグラム差信号を第1のバッファ手段から読み出し、読み出した輝度ヒストグラム差信号の平均値

5 を求める。この平均値が第8の閾値Th8以上になるとき、検出された目間候補の左側にある目候補の目候補信号の右下の座標と、検出された目間候補の右側にある目候補の目候補信号の左上の座標を、目間領域信号118として出力する。

手順211では、顔決定手段120が、目間領域信号118の表す座標（目の間の領域の右下の座標と左上の座標）を基準にして、あらかじめ設定しておいた

10 変換式に従って顔の領域の座標を決定し、顔領域信号121として出力する。

目の間の領域の左上の座標を(xli,yti)とし、右下の座標を(xri,ybi)とすると、例えば、顔の領域の左上の座標(xlf,ytf)と右下の座標(xrf,ybf)を

$$\begin{aligned} xlf &= xli - (xri - xli) \\ xrf &= xri + (xri - xli) \\ 15 \quad ytf &= yti \\ ybf &= yti + (xrf - xlf) \end{aligned}$$

とする。

本実施の形態で人物の顔を検出した例として、図3（b）に目候補領域303、304と目の間の候補領域308を示し、図3（c）に目の間の候補領域308から求めた目の間の領域の左上の座標点301と右下の座標点302と、顔の領域の左上の座標点305と右下の座標点306を示す。

20

以上の説明では、画像全体を小領域に分割しているが、画像の一部（目の間の領域が存在すると予想される領域）を小領域に分割してもよい。例えば、テレビ放送のフレーム画像から目の間の領域を検出する場合には、顔が画像の上半分に

25 存在することが多いので、画像の上半分を小領域に分割し、この領域から目の間の領域を検出すればよい。

また、目候補を検出する手順206、目間候補を検出する手順207および手順208は順次処理するように説明したが、並列または手順206～手順208をどんな順番で実行してもよい。

手順202では、一般に知られている他の方法を用いてエッジ強度を計算してもよい。また、輝度差の二乗の総和ではなく、輝度差の絶対値の総和をエッジ特徴量にしてもよい。

同様に、手順204では、一般に知られている他の方法を用いて輝度ヒストグラム5の差を計算してもよい。また、頻度差 $[H(i,j,bin)-H(i-1,j,bin)]$ の二乗の総和ではなく、頻度差の絶対値の総和を輝度ヒストグラム差信号として出力してもよい。エッジ強度や輝度ヒストグラムの差を利用せずに、他の特徴量を利用して直線や曲線の有無を調べて、目候補や目間候補を検出してもよい。

10 眼鏡をかけた顔の目の間の候補は、他の特徴を利用したり、複数の特徴を組み合わせ検出してもよい。本実施の形態の手順208では、額部分（眼鏡の上側）の特徴を利用しているが、例えば鼻部分（眼鏡の下側）で線が少ない（各画素のエッジ強度が非常に小さい）ことを利用してもよい。

また、眼鏡をかけている顔の両目の間の輝度分布の特徴（例えば、横方向に隣接する小領域の輝度ヒストグラムの差が大きくなるという特徴）を利用してもよい。15

目の間の領域から顔の領域を決定する方法として、他の方法を利用してもよい。例えば、目の間の領域の周辺に顔の候補領域を設定し、顔の候補領域内で肌色になる画素を肌色画素として取り出してから、あらかじめ設定した割合以上の肌色画素を含むような矩形領域を求めて、顔の領域としてもよい。

20 本実施の形態によれば、両目と目の間の特徴を利用して目の間の領域を検出するので、口の動きを含む顔の表情やヒゲの有無や髪型による顔の変動に強く、人物の顔の領域を検出することができる。

目の領域では、目や瞼の動きを含む顔の変動や眼鏡の有無や髪型によらず垂直方向のエッジ強度が大きい値になり、ほぼ同じ値の輝度をもつ画素が横方向に隣接するので、目や瞼の動きを含む顔の変動や眼鏡の有無や髪型によらず、目の領域25の候補を検出できる。また、眼鏡をかけたときの目の間の特徴と眼鏡をかけないときの目の間の特徴を両方利用して目の間の候補を検出することから、眼鏡の有無によらず目の間の領域の候補を検出できる。さらに、顔の表情が変動しても目の間がほとんど変動しないことから、顔の表情の変動によらず、目の間の領域

の候補を検出できる。そして、本実施の形態で利用している目や目の間の特徴は、頭全体が動いて顔の向きや顔の傾きが変動しても保持されるので、頭全体の動きによらず、目の間の候補を検出できる。以上のことから、本実施の形態の人物の顔の検出装置は、目や瞼や頭全体の動きを含む顔の変動や眼鏡の有無や髪型によらず、人間の顔の領域を検出できる。

さらに、最初に小領域に分割してから小領域が目や目の間であるかどうかを判定するので、複数の画素が連結領域かどうかを判定するような多くの計算量を要する処理が不要であり、少ない計算量で人物の顔の領域を検出できる。

人物の顔を検出する技術を使って動画像を処理する装置やシステムでは、少ない計算量で高速に人物の顔の領域を検出しなければならないことが多い。本実施の形態によれば、輝度極小点と輝度変曲点に挟まれる領域の連結領域を調べて目や口の候補領域にする処理のような、多くの計算量が必要な処理を実行しないので、高速に人物の顔の領域を検出することができる。また、人物の顔の領域を検出することで、画像中の顔の数（人数）を簡単に推定できる。

図2に示した処理手順を人物の顔の検出プログラムとして記憶媒体に記憶することにより、各種コンピュータにインストールして人物の顔の検出方法を実行し、画像から人物の顔の領域を検出することが可能となる。

（実施の形態2）

実施の形態2では、実施の形態1に示した目間領域検出手段101を2つ組み合わせ、より高精度な人物の顔の検出装置を提供する。

図4は、本発明の実施の形態2における人物の顔の検出装置の構成図である。図4において、人物の顔の検出装置401には、入力画像信号402が部分画像発生手段403に入力される。

部分画像発生手段403は、入力された入力画像信号402の画像全体から、図5に示す領域501の画像と領域502の画像を取り出し、第1の部分画像信号404と第2の部分画像信号405を出力する。

第1の目間領域検出手段406は、第1の部分画像信号404を読み込んで目の間の領域を検出し、目の間の領域の座標を第1の目間領域候補信号407として出力する。

第2の目間領域検出手段408は、第2の部分画像信号405を読み込んで目の間の領域を検出し、目の間の領域の座標を第2の目間領域候補信号409として出力する。

バッファ手段410は、第1の目間領域候補信号407と第2の目間領域候補信号409を記憶する。

重複判定手段411は、バッファ手段410から第1の目間領域候補信号と第2の目間領域候補信号を読み込んで、2つの検出された目の間の領域が重複する部分の面積を判定して、目の間の領域の座標を目間領域信号412として出力する。

顔決定手段413は、目間領域信号412を読み込んで、顔の領域の座標を表す顔領域信号414を出力する。

目間領域検出手段406、408は、実施の形態1に示した目間領域検出手段101と同じ構成にすればよい。また、実施の形態1に示した人物の顔の検出装置における小領域の高さを V_{slit} とし、座標系の原点(0,0)を画像の左上端の点にしたとき、図5に示した領域501の左上端の座標を例えば(1,1)とし、領域502の左上端の座標を(2, $V_{slit}/2+1$)にすればよい。これにより、第1の目間領域検出手段406におけるそれぞれの小領域は、第2の目間領域検出手段408におけるどこかの小領域と必ず重なることになる。なお、領域501を分割して得られた小領域の集合を第1の小領域群とし、領域502を分割して得られた小領域の集合を第2の小領域群とする。

以下、人物の顔の検出装置の動作について、図6のフローチャートを用いて説明する。

手順601では、部分画像発生手段403が、画像入力信号402によって入力された画像全体から、図5に示す領域501と領域502を取り出し、領域501の画像を第1の部分画像信号404として出力し、領域502の画像を第2の部分画像信号405として出力する。

手順602では、第1の目間領域検出手段406が、第1の部分画像信号404を読み込んで目の間の領域を検出し、目の間の領域の座標を第1の目間領域候補信号407として出力する。例えば、実施の形態1の手順201から手順21

0に示した動作を実行することで、目の間の領域を検出することができる。

手順603では、第2の目間領域検出手段408が、第2の部分画像信号405を読み込んで目の間の候補領域を検出し、目の間の領域の座標を第2の目間領域候補信号409として出力する。

- 5 手順604では、バッファ手段410が、第1の目間領域候補信号407と第2の目間領域候補信号409を記憶する。図7に示した人物の顔の画像から本装置で人物の顔を検出したとき、第1の目間領域検出手段で検出された目の間の領域の左上の座標と右下の座標は点701、702になり、第2の目間領域検出手段で検出された目の間の領域の左上の座標と右下の座標は点703、704になる。

- 10 手順605では、重複判定手段411が、バッファ手段410から第1の目間領域候補信号と第2の目間領域候補信号を読み込み、第1の目間領域検出手段406で検出された目の間の領域の左上の座標(xl1,yt1)および右下の座標(xr1,yb1)と、第2の目間領域検出手段408で検出された目の間の領域の左上の座標(x12,yt2)および右下の座標(xr2,yb2)を求め、2つの目の間の領域に重複する部分が存在するかを調べる。

すべての目間領域候補信号の組み合わせに対して手順607以降を実行した場合には、人物の顔の検出処理を終了する（手順606）。

- 20 2つの目の間の領域に重複する部分が存在しない場合には、手順605に戻る（手順607）。

2つの目の間の領域に重複する部分が存在する場合には、重複する部分の左上の座標(xlo,yto)および右下の座標(xro,ybo)を求める（手順608）。

- 25 手順609では、2つの座標(xl1,yt1)、(xr1,yb1)で囲まれる領域（第1の目の間の領域）の面積S1と、2つの座標(xl2,yt2)、(xr2,yb2)で囲まれる領域（第2の目の間の領域）の面積S2と、2つの座標(xlo,yto)、(xro,ybo)で囲まれる領域（重複する部分の領域）との面積Soを求め、その割合が閾値Th以上になることを表す式

$$So/S1 > Th$$

$$So/S2 > Th$$

が同時に成立するときには、重複する部分の左上の座標(xlo,yto)および右下の座標(xro,ybo)を目間領域信号412として出力する。

- 手順610では、顔決定手段413が、目間領域信号412の表す座標を基準にして、あらかじめ設定しておいた変換式に従って顔の領域の座標を決定し、顔領域信号414として出力する。手順610を実行した後は手順605に戻る。

以上の説明では、第1の小領域群と第2の小領域群の間で、分割される領域が異なるようにしているが、小領域の大きさが異なるようにしてもよい。

- 本実施の形態によれば、2つの検出された目の間の領域が重複するとき、両方とも画像中の目の間を含んでいれば2つの領域の重複する部分の面積が大きくなるので、顔以外の背景物が多く映っている場合を含めて、人物の顔の領域の誤検出を抑制し、高精度に検出することができる。

また、少ない計算量で候補領域の重なりを判定するので、高速に実行することができる。

- さらに、目の間の候補領域の重なりを調べるだけなので、口の動きを含む顔の表情やヒゲの有無や髪型や頭全体の動きによる顔の変動に強く、人物の顔の領域の誤検出を抑制し、高精度に検出することができる。

(実施の形態3)

- 実施の形態3では、実施の形態1に示した目間領域検出手段と、目の間の領域の周辺を調べる手段を組み合わせ、より高精度な人物の顔の検出装置を提供する。

図8は本発明の実施の形態3における人物の顔の検出装置の構成図である。図8において、人物の顔の検出装置821には、入力画像信号802が入力され、入力画像信号がそのまま目間領域検出手段801の目間候補設定手段803に入力される。

- 目間候補設定手段803は、入力画像信号802を読み込んで目の間の候補領域を設定し、目の間の候補領域の座標を目間領域候補信号804として出力する。

目候補設定手段805は、目間領域候補信号804を読み込んで、目の間の候補領域の左右に目の候補領域を設定し、左目領域候補信号806と右目領域候補

信号 807 を出力する。

バッファ手段 808 は、目間領域候補信号 804 と左目領域候補信号 806 と右目領域候補信号 807 を記憶する。

部分画像発生手段 809 は、入力画像信号 802 によって入力された画像全体
5 から、目間領域候補信号と左目領域候補信号と右目領域候補信号の表す領域の画像を取り出し、目間画像信号 810 と左目画像信号 811 と右目画像信号 812 を出力する。

輝度比較手段 813 は、目間画像信号 810 と左目画像信号 811 と右目画像信号 812 を読み込んで、目の間の候補領域の平均輝度と左目の候補領域の平均
10 輝度と右目の候補領域の平均輝度を比較して、第 1 の判定結果信号 814 を出力する。

色判定手段 814 は、目間画像信号 810 と左目画像信号 811 と右目画像信号 812 を読み込んで、目の間の候補領域の平均色と左目の候補領域の平均色と右目の候補領域の平均色を調べて、第 2 の判定結果信号 816 を出力する。

15 目間決定手段 817 は、第 1 の判定結果信号 815 と第 2 の判定結果信号 816 を読み込んで、目の間の領域の座標を目間領域信号 818 として出力する。

顔決定手段 819 は、目間領域信号 818 を読み込んで、顔の領域の座標を表す顔領域信号 820 を出力する。

特徴判定手段 822 は、輝度比較手段 813 と色判定手段 814 から構成され
20 る。目間候補設定手段 803 は、実施の形態 1 に示した目間領域検出手段と同じ構成にすればよい。また、実施の形態 2 に示した人物の顔の検出装置から顔決定手段を除いた部分と同じ構成にして、目間領域信号 412 を出力するようにしてもよい。

以下、人物の顔の検出装置の動作について、図 9 のフローチャートを用いて説
25 明する。

手順 901 では、目間候補設定手段 803 が、入力画像信号 802 を読み込んで目の間の領域を検出し、目の間の候補領域にする。そして、目の間の候補領域の座標を目間領域候補信号 804 として出力する。例えば、実施の形態 1 に示した手順 201 から手順 211 に示す動作を実行することで、目の間の領域を検出

することができる。また、実施の形態2に示した手順601から手順609を実行してもよい。

手順902では、目候補設定手段805が、目間領域候補信号804から目の間の候補領域の左上の座標(x_{li}, y_{ti})と右下の座標(x_{ri}, y_{bi})を読み取る。そして、左目の候補領域の左上の座標(x_{ll}, y_{tl})と右下の座標(x_{rl}, y_{bl})と、右目の候補領域の左上の座標(x_{lr}, y_{tr})と右下の座標(x_{rr}, y_{br})を設定し、左目の候補領域の座標を左目領域候補信号806として出力し、右目の候補領域の座標を右目領域候補信号807として出力する。図10に目の間の候補領域の左上の座標1001と右下の座標1002と、左目の候補領域の左上の座標1003と右下の座標1004と、右目の候補領域の左上の座標1005と右下の座標1006の例を示す。 x_{ll} 、 y_{tl} 、 x_{rl} 、 y_{bl} 、 x_{lr} 、 y_{tr} 、 x_{rr} 、 y_{br} の値は、次式で決定する。

$$\begin{aligned} x_{ll} &= x_{ri} \\ y_{tl} &= y_{ti} \\ x_{rl} &= x_{ll} + (x_{ri} - x_{li}) / 2 \\ y_{bl} &= y_{bi} \\ x_{lr} &= x_{rr} - (x_{ri} - x_{li}) / 2 \\ y_{tr} &= y_{ti} \\ x_{rr} &= x_{li} \\ y_{br} &= y_{bi} \end{aligned}$$

手順903では、バッファ手段808が、目間領域候補信号804と左目領域候補信号806と右目領域候補信号807を記憶する。

手順904では、部分画像発生手段809が、入力画像信号802によって入力された画像全体から、目間領域候補信号と左目領域候補信号と右目領域候補信号の表す領域の画像を取り出す。そして、目の間の候補領域の画像を目間画像信号810として出力し、左目の候補領域の画像を左目画像信号811として出力し、右目の候補領域の画像を右目画像信号812として出力する。

手順905では、輝度比較手段813が、目間画像信号810と左目画像信号811と右目画像信号812を読み込んで、目の間の候補領域における輝度の平均値 Y_i と左目の候補領域における輝度の平均値 Y_l と右目の候補領域における輝度

の平均値 Y_r を求める。そして、 Y_i が Y_l よりも大きく、かつ、 Y_i が Y_r よりも大きいとき、第1の判定結果信号815を1（目の間の可能性あり）にする。そうでなければ、第1の判定結果信号815を0（目の間の可能性なし）にする。

手順906では、色判定手段814が、目間画像信号810と左目画像信号811と右目画像信号812を読み込んで、目の間の候補領域における色の平均(R_i, G_i, B_i)と左目の候補領域における色の平均(R_l, G_l, B_l)と右目の候補領域における色の平均(R_r, G_r, B_r)を求める。ただし、 R_i 、 R_l 、 R_r は色の赤成分を表し、 G_i 、 G_l 、 G_r は色の緑成分を表し、 B_i 、 B_l 、 B_r は色の青成分を表す。肌色は赤成分が多く青成分が少ないので、

$$R_i > G_i$$

$$G_i > B_i$$

$$R_l > G_l$$

$$G_l > B_l$$

$$R_r > G_r$$

$$G_r > B_r$$

を同時に満足するとき、第2の判定結果信号816を1（目の間の可能性あり）にする。そうでないとき、第2の判定結果信号816を0（目の間の可能性なし）にする。

手順907では、第1の判定結果信号815と第2の判定結果信号816が同時に1になるとき、目間決定手段817が目目の間の領域の座標を目間領域信号417として出力する。

手順908では、顔決定手段819が、目間領域信号817の表す座標を基準にして、あらかじめ設定しておいた変換式に従って顔の領域の座標を決定し、顔領域信号820として出力する。

以上の説明では、左目の候補領域の幅と右目の候補領域の幅を目目の間の候補領域の半分にしているが、画像によって左右の目の候補領域の幅と位置を変えてもよい。例えば、手順902において、目の間の候補領域の右側と左側に目の探索領域を設定して、探索領域内の輝度特徴や形状特徴を利用して左目の候補領域と右目の候補領域の幅と位置を決定してもよい。

また、手順906において、色の赤成分と青成分と緑成分の大小関係だけを用いて第2の判定結果信号の値を決定しているが、他の方法を用いて決定してもよい。例えば、あらかじめ色空間上における肌色領域を調べておき、目の間の候補領域や左目の候補領域や右目の候補領域における色の平均が肌色領域内に含まれるかどうかを判定して、第2の判定結果信号の値を決定してもよい。また、左目の候補領域における色の平均と右目の候補領域における色の平均が近い色であるかどうかを判定して、第2の判定結果信号の値を決定してもよい。左目の候補領域の上半分における色の平均と目の間の候補領域の下半分における色の平均が近い色であるかどうかを判定して第2の判定結果信号の値を決定してもよい。

10 輝度比較手段813による判定結果と色判定手段814による判定結果の両方を用いて目の間の領域を決定しているが、片方の判定結果だけを用いてもよい。また、ニューロや統計的手法を実行する他の手段を併用して目の間の領域を決定してもよい。また、右目と左目の候補領域を検出してから、目の間の候補領域を決定してもよい。

15 本実施の形態によれば、顔の表情の変動や頭全体の動きによらず目の間の領域は目の領域に比べて輝度値が大きいので、顔以外の背景物が多く映っている場合を含めて、目や口の動きを含む顔の表情やヒゲの有無や頭全体の動きによる顔の変動に強く、少ない計算量で人物の顔の領域を検出できる。

また、目の周辺や目の間の色を平均すると肌の色に近い色になるので、顔以外の背景物が多く映っている場合を含めて、人物の顔の領域の誤検出を抑制し、高精度に検出することができる。

(実施の形態4)

図11は、本発明の実施の形態4における人物の顔の検出装置の構成図である。図11において、人物の顔の検出装置1115には、入力画像信号1102が25 入力され、この入力画像信号がそのまま目検出手段1101に入力される。

目間領域検出手段1103は、入力画像信号1102を読み込んで目の間の領域を検出し、目の間の領域の座標を目間領域信号1104として出力する。

目候補設定手段1105は、目間領域信号1104を読み込んで、目の間の領域の左右に目の候補領域を設定し、左目領域候補信号1106と右目領域候補信

号1107を出力する。

- 目決定手段1108は、入力画像信号1102によって入力された画像全体から、左目領域候補信号1106と右目領域候補信号1107の表す領域の画像を取り出し、それぞれの画像の輝度特徴または形状特徴を調べて右目の位置と左目
- 5 の位置を決定し、左目位置信号1109と右目位置信号1110を出力する。

顔決定手段1111は、左目位置信号1109と右目位置信号1110を読み込んで、顔の領域の座標を表す顔領域信号1112を出力する。

- 図12は、実施の形態4における目決定手段1108の構成図である。目決定手段1108には、入力画像信号1102と左目領域候補信号1106と右目領域候補信号1107が入力される。
- 10

バッファ手段1201は、左目領域候補信号1106と右目領域候補信号1107を記憶する。

テンプレート読取手段1202は、あらかじめ用意した標準的な目のテンプレートを読み込み、標準目画像信号1203として出力する。

- 部分画像発生手段1204は、両目の候補領域内からテンプレートと同じサイズの画像を切り出し、候補画像信号1205として出力する。
- 15

部分画像発生手段1204で切り出す領域は、領域指定信号1206によって指定される。

- 第1のマッチング手段1207は、標準目画像信号1203と候補画像信号1205を比較して左目の位置を決定し、左目位置信号1208を出力する。
- 20

第2のマッチング手段1209は、標準目画像信号1203と候補画像信号1205を比較して右目の位置を決定し、右目位置信号1210を出力する。

目間領域検出手段1103は、例えば実施の形態3に示した目間領域検出手段801と同じ構成にすればよい。

- 以下、人物の顔の検出装置の動作について、図13のフローチャートを用いて説明する。
- 25

手順1301では、目間領域検出手段1103が、入力画像信号1102を読み込んで目の間の領域を検出し、目の間の領域の座標を目間領域信号1104として出力する。例えば、実施の形態3に示した動作を実行することで、目の間の

領域を検出することができる。

手順 1 3 0 2 では、目候補設定手段 1 1 0 5 が、目間領域信号 1 1 0 4 から目の間の領域の左上の座標(xli,yti)と右下の座標(xri,ybi)を読み取る。そして、左目の候補領域の左上の座標(xll,ytl)と右下の座標(xrl,ybl)と、右目の候補領域の左上の座標(xlr,ytr)と右下の座標(xrr,ybr)を設定し、左目の候補領域の座標を左目領域候補信号 1 1 0 6 として出力し、右目の候補領域の座標を右目領域候補信号 1 1 0 7 として出力する。目の間の領域の左上の座標 1 4 0 1 と右下の座標 1 4 0 2 と、左目の候補領域の左上の座標 1 4 0 3 と右下の座標 1 4 0 4 と、右目の候補領域の左上の座標 1 4 0 5 と右下の座標 1 4 0 6 の例を図 1 4 に示す。xll、ytl、xrl、ybl、xlr、ytr、xrr、ybrの値は、次式で決定する。

$$\begin{aligned} xll &= xri - (xri - xli) / 3 \\ ytl &= yti - (ybi - yti) \\ xrl &= xll + (xri - xli) \\ ybl &= ybi + (ybi - yti) \\ xlr &= xrr - (xri - xli) \\ ytr &= yti - (ybi - yti) \\ xrr &= xli + (xri - xli) / 3 \\ ybr &= ybi + (ybi - yti) \end{aligned}$$

手順 1 3 0 3 では、バッファ手段 1 2 0 1 が、左目領域候補信号 1 1 0 6 と右目領域候補信号 1 1 0 7 を記憶する。

手順 1 3 0 4 では、テンプレート読取手段 1 2 0 2 が、あらかじめ用意した標準的な目のテンプレートを読み込み、標準目画像信号 1 2 0 3 として出力する。

手順 1 3 0 5 では、第 1 のマッチング手段 1 2 0 7 がバッファ手段 1 2 0 1 から左目領域候補信号を読み取り、左目の候補領域内にテンプレートと同じサイズの領域を候補画像領域として設定する。多数の候補画像領域が設定されることになる。n 番目 (n は 1 以上の整数) の候補画像領域の座標は、n 番目の領域指定信号 1 2 0 6 として出力される。

手順 1 3 0 6 では、部分画像発生手段 1 2 0 4 が、それぞれの候補画像領域の画像を切り出し、候補画像信号 1 2 0 5 として出力する。

手順1307では、第1のマッチング手段1207が、標準目画像信号1203と候補画像信号1205を比較して、標準目画像信号との一致度がもっとも大きい候補画像信号を求める。標準目画像信号との一致度が最も大きい候補画像信号の番号を調べて、その番号の候補画像領域の座標を左目位置信号1208として出力する。

手順1308では、第2のマッチング手段1209がバッファ手段1201から右目領域候補信号を読み取り、右目の候補領域内にテンプレートと同じサイズの領域を候補画像領域として設定する。多数の候補画像領域が設定されることになる。n番目（nは1以上の整数）の候補画像領域の座標は、n番目の領域指定信号1206として出力される。

手順1309では、部分画像発生手段1204が、それぞれの候補画像領域の画像を切り出し、候補画像信号1205として出力する。

手順1310では、第2のマッチング手段1209が、標準目画像信号1203と候補画像信号1205を比較して、標準目画像信号との一致度がもっとも大きい候補画像信号を求める。標準目画像信号との一致度が最も大きい候補画像信号の番号を調べて、その番号の候補画像領域の座標を右目位置信号1210として出力する。

手順1311では、顔決定手段1111が、左目位置信号1109から左目の中心の座標を調べ、右目位置信号1110から右目の中心の座標を調べ、それぞれの座標を基準にして、あらかじめ設定しておいた変換式に従って顔の領域の座標を決定し、顔領域信号1112として出力する。左目の中心の座標を(xl,yl)とし、右目の中心の座標を(xr,yr)とすると、例えば、顔の領域の左上の座標(xl_f,yl_f)と右下の座標(xr_f,yr_f)を

$$xl_f = xr - (xl - xr) / 2$$

$$xr_f = xl + (xl - xr) / 2$$

$$yl_f = (yl + yr) / 2 - (xr_f - xl_f) * 0.2$$

$$yr_f = (yl + yr) / 2 + (xr_f - xl_f) * 0.8$$

とする。

以上の説明では、両目の候補領域の画像に対するテンプレートマッチングを用

いて目の位置を決定しているが、目決定手段1108の構成を変更し、テンプレートマッチング以外の方法を用いて目の位置を決定してもよい。例えば、両目の候補領域内におけるエッジ方向の対称性を用いて目の位置を決定してもよいし、先験知識を用いて目の位置を決定してもよい。また、左目の位置を決定してから右目の位置を決定してもよい。

本実施の形態によれば、目の間は顔の表情や頭全体の動きによる変動が少なく口やヒゲと離れているので、目や口の動きを含む顔の表情やヒゲの有無や頭全体の動きによる顔の変動に強く、人物の顔や目の位置を検出できる。

また、目の周辺だけに限定すれば、目の輝度分布と目以外の部分の輝度分布が大きく異なるので、目決定手段によって右目と左目の位置を検出できる。

(実施の形態5)

図15は、本発明の実施の形態5における人物の顔の検出装置の構成図である。図15において、人物の顔の検出装置1501には、入力画像信号1502が入力される。目間領域検出手段1503は、入力画像信号1502を読み込んで目の間の領域を検出し、目の間の領域の座標を目間領域信号1504として出力する。

顔決定手段1505は、目間領域信号1504と入力画像信号1502を読み込んで、目の間の領域の周辺から、目の間の領域の平均色に近い色の領域を決定し、顔の領域の座標を表す顔領域信号1506を出力する。

目間領域検出手段1503は、例えば実施の形態3に示した目間領域検出手段と同じ構成にすればよい。

以下、人物の顔の検出装置の動作について、図16のフローチャートを用いて説明する。

手順1601では、目間領域検出手段1503が、入力画像信号1502を読み込んで目の間の領域を検出し、目の間の領域の座標を目間領域信号1504として出力する。例えば、実施の形態3に示した動作を実行することで、目の間の領域を検出することができる。

手順1602では、顔決定手段1505が、目の間の領域内における画素の色の平均 C_m を計算し、色空間内で平均色 C_m を中心とする球状の領域に含まれる色を

肌色とする。

手順 1 6 0 3 では、顔決定手段 1 5 0 5 が、目の間の領域の周辺に顔の候補領域を設定し、顔の候補領域内で肌色になる画素を肌色画素として取り出してから、あらかじめ設定した割合以上の肌色画素を含むような矩形領域を求めて、顔の領域とする。そして、顔の領域の左上の座標と右下の座標を顔領域信号 1 5 0 6 として出力する。

以上の説明では、目の間の領域に含まれる色を用いて顔の領域を決定しているが、他の方法を用いて顔の領域を決定してもよい。例えば、実施の形態 1 ～ 4 に示したように、あらかじめ設定しておいた変換式に従って顔の領域の座標を決定してもよい。

本実施の形態によれば、目の間は顔の表情や頭全体の動きによる変動が少なく口やヒゲと離れているので、目や口の動きを含む顔の表情やヒゲの有無や頭全体の動きによる顔の変動に強く、顔の領域を検出できる。また、照明条件や個人差によらず目の間の領域の色と顔の肌の色は近いので、照明条件や個人差の変動に強く、少ない計算量で顔の領域を検出できる。

(実施の形態 6)

図 1 7 は、本発明の実施の形態 6 における動画像検索装置の構成図である。図 1 7 において、動画像蓄積手段 1 7 0 1 は、動画像ファイルを蓄積する。

画像発生手段 1 7 0 2 は、動画像選択信号 1 7 2 3 で指定された動画像ファイルの動画像を表す動画像信号 1 7 0 3 を、動画像蓄積手段 1 7 0 1 から読み込んで、フレーム画像を表す画像信号 1 7 0 4 とフレーム画像の時刻を表す時刻信号 1 7 1 9 を出力する。

顔特徴検出手段 1 7 0 5 は、画像信号 1 7 0 4 を読み込んで、目の間の領域を検出して目の間の領域の座標を目間領域信号 1 7 0 6 として出力する。

顔決定手段 1 7 0 7 は、目間領域信号 1 7 0 6 と画像信号 1 7 0 4 を読み込んで、目の間の領域の周辺から、目の間の領域の平均色に近い色の領域を決定し、顔の領域の座標を表す顔領域信号 1 7 0 8 を出力する。

情報検出手段 1 7 0 9 は、顔領域信号 1 7 0 8 と画像信号 1 7 0 4 を読み込んで、顔の領域の画像を取り出し、性別、年齢を求め顔情報信号 1 7 1 0 とし、顔

画像信号1724および顔情報信号1710を出力する。

顔情報蓄積手段1711は、顔画像信号1724と顔情報信号1710と時刻信号1719とを動画像情報ファイルとして蓄積する。

5 入力手段1712は、動画像ファイルを選択したり、検索の条件式を入力したり、顔画像を選択したりする。

検索手段1713は、入力手段からの検索条件信号1714を読み取って、顔情報蓄積手段1711に蓄積された動画像情報ファイルの中から検索の条件式に適合する情報を取り出し、検索結果信号1715として出力する。

10 検索結果整形手段1716は、検索結果信号1715を記憶したり、検索結果をユーザーにわかりやすい形に整形して画面イメージ信号1720として出力したり、顔画像の選択結果を表す画像選択信号を読み込んで動画像を再生するための再生制御信号1722を出力する。

表示手段1717は、画面イメージ信号1720を読み取り、動画像や検索結果などを表示する。

15 再生手段1718は、動画像蓄積手段1701に蓄積された動画像を表示手段1717に表示するために、動画像信号を画面イメージ信号1720に変換する。

20 顔特徴検出手段1705は、例えば実施の形態3に示した目間検出手段と同じ構成にすればよい。また、動画像ファイルの形式としては、圧縮されたMPEG1形式やMPEG2形式やDV形式や非圧縮形式などを利用することができる。動画像蓄積手段1701と顔情報蓄積手段1711はハードディスク、DVD-RAM、PDなどの記憶媒体によって実現できる。画像発生手段1702、顔特徴検出手段1705、顔決定手段1707、情報検出手段1709、検索手段1713、検索結果整形手段1716、再生手段1718は、パソコンのCPUによって実現できる。表示手段1717はモニタ等の表示デバイスによって実現できる。入力手段1712はキーボードと、マウス等のポインティングデバイスによって実現できる。

動画像検索装置では、あらかじめ動画像情報ファイルを作成しておいてから、検索手段を使って顔画像を検索する。

最初に、動画像検索装置を用いて動画像情報ファイルを作成する動作について、図18のフローチャートを用いて説明する。

手順1801では、動画像蓄積手段1701に蓄積されている動画像ファイルの一つを選択する。選択結果は、動画像選択信号として画像発生手段1702に入力される。

手順1802では、画像発生手段1702が、動画像蓄積手段1701から動画像信号1703を読み込んで、手順1801で選択した動画像ファイルの1枚のフレーム画像を取り出し、画像信号1704として出力する。

手順1803では、顔特徴検出手段1705が、画像信号1704を読み込んで目の間の領域を検出し、目の間の領域の座標を目間領域信号1706として出力する。例えば、実施の形態3に示した動作を実行することで、目の間の領域を検出することができる。

手順1804では、実施の形態5の手順1602と同様に、顔決定手段1707が、目の間の領域内における画素の色の平均 C_m を計算し、色空間内で平均色 C_m を中心とする球状の領域に含まれる色を肌色とする。

手順1805では、実施の形態5の手順1603と同様に、顔決定手段1707が、目の間の領域の周辺に顔の候補領域を設定し、顔の候補領域内で肌色になる画素を肌色画素として取り出してから、あらかじめ設定した割合以上の肌色画素を含むような矩形領域を求めて、顔の領域とする。そして、顔の領域の左上の座標と右下の座標を顔領域信号1708として出力する。

手順1806では、情報検出手段1709が、男女識別関数 $g1$ と年齢識別関数 $g2$ を実行する。男女識別関数 $g1$ は、顔領域信号の値と、画像信号の表すフレーム画像を引数として、男女識別し、識別結果 $res1$ を返す。年齢識別関数 $g2$ は、顔領域信号の値と、画像信号の現すフレーム画像を引数として、年齢識別し、識別結果 $res2$ を返す。

情報検出手段1709は、顔の領域の左上の座標と右下の座標と識別結果 $res1$ 、 $res2$ をまとめて顔情報信号1710として出力し、さらに、画像信号の表すフレーム画像中から顔領域信号の表す顔の領域の画像を切り出して顔画像信号1724として出力する。

- 男女識別関数 $g1$ の実現方法としては、予め性別が既知の顔画像を多数収集し、この収集した顔画像を元に判別分析等の統計的手法により実現することができる。同様に年齢識別関数 $g2$ は、予め年代が既知の顔画像を多数収集し、この収集した顔画像を元に判別分析等の統計的手法により実現することができる。なお、ニューラルネットワークの学習により、識別関数を実現することもできる。

- 手順1807では、顔情報蓄積手段1711が、時刻信号1719の表すフレーム画像の時刻と、顔情報信号1710の表す情報を動画像情報ファイルとして蓄積する。また、現在のフレーム画像に近い時刻のフレーム画像（例えば0.5秒前のフレーム画像）から検出された顔の位置と、現在のフレーム画像から検出された顔の位置とが近い場合を除いて、顔画像信号の表す顔の領域の画像を顔画像ファイルとして蓄積する。

- 手順1808では、別のフレーム画像から顔に関する情報を取り出すため、手順1802に戻る。ただし、手順1802から手順1807までの処理を、動画像の末尾まで実行した場合は、手順1802に戻らずに動画像情報ファイルと顔画像ファイルを作成する処理を終了する。また、顔は少なくとも数秒以上画面上に存在するので、手順1802ですべてのフレーム画像を取り出す必要はない。例えば0.5秒おきのフレーム画像を取り出してもすべての登場人物の顔を検出することが可能である。

- 次に、検索手段を使って顔画像を検索する動作について、図19のフローチャートを用いて説明する。

手順1901では、ユーザーが入力手段1712を用いて検索の条件式を入力する。

- 手順1902では、検索手段1713が検索の条件式を読み取って、顔情報蓄積手段1711に蓄積された動画像情報ファイルの中から検索の条件式に適合する情報を取り出す。そして、顔に関するデータ（性別、年齢、顔の存在するフレーム画像の時刻）と顔の領域の画像（顔画像）を検索結果信号として出力する。例えば、検索の条件式が「男性」であるとき、男性の顔であると推定されたすべての顔画像と、それぞれの顔画像に関連するデータを、検索結果信号として出力する。

手順1903では、検索結果整形手段1716が、検索結果信号1715を記憶する。

手順1904では、検索結果整形手段1716がユーザーにわかりやすい形に整形して、表示手段に表示する。例えば、検索結果信号に含まれていた顔画像と

5 、性別、年齢を表示する。

手順1905では、ユーザーが入力手段1712を用いて顔画像を一つ選択する。選択した結果は、画像選択信号として検索結果整形手段1716に入力される。

10 手順1906では、検索結果整形手段1716が、再生手段1718を制御しながら、手順1905で選択された顔画像の人物のシーンを次々に再生する。

以上の説明では、顔特徴検出手段1705で目の間の領域を検出しているが、鼻、耳、口などの他の領域を検出し、その領域を含むように顔決定手段で顔の領域を決定してもよい。

15 本実施の形態によれば、目の間は顔の表情や頭全体の動きによる変動が少なく口やヒゲと離れているので、目や口の動きを含む顔の表情やヒゲの有無や頭全体の動きによる顔の変動に強く、顔の領域を検出できる。また、人物や顔をキーにして動画像を検索したり、検索のための情報を作成して蓄積することができる。

(実施の形態7)

20 図20は、本発明の実施の形態7における顔認識装置の構成図である。顔認識装置は、予め多くの人の身分証がイメージスキャナ等により顔画像をデータベースに登録しておき、ビデオカメラで撮像した顔画像の人物の該当者がデータベースの中にいるか、あるいは登録された身分証のどれに一番類似しているかを認識するものである。

25 図20において、ビデオカメラ2001は、部屋の出入り口のような監視対象領域の動画像を動画像信号2002として出力する。画像発生手段2003は、動画像信号2002を読み込んで、フレーム画像を表す画像信号2004を出力する。

目間領域検出手段2005は、画像信号2004を読み込んで、目の間の領域を検出して目の間の領域の座標を目間領域信号2006として出力する。

手順 2 1 0 2では、実施の形態 4 の手順 1 3 0 1と同様に、目間領域検出手段

2005が、画像信号2004を読み込んで目の間の領域を検出し、目の間の領域の座標を目間領域信号2006として出力する。

手順2103では、実施の形態4の手順1302と同様に、目候補設定手段2007が、目間領域信号2006から目の間の領域の左上の座標(xli,yti)と右下の座標(xri,ybi)を読み取る。そして、左目の候補領域の左上の座標(xll,ytl)と右下の座標(xrl,ybl)と、右目の候補領域の左上の座標(xlr,ytr)と右下の座標(xrr,ybr)を設定し、左目の候補領域の座標を左目領域候補信号2008として出力し、右目の候補領域の座標を右目領域候補信号2009として出力する。

手順2104では、目決定手段2010が、画像信号2004によって入力された画像全体から、左目領域候補信号2008と右目領域候補信号2009の表す領域の画像を取り出し、それぞれの画像の輝度特徴または形状特徴を調べて右目の位置と左目の位置を決定し、左目位置信号2011と右目位置信号2012を出力する。例えば、実施の形態4の手順1303～手順1310と同様の動作を実行すればよい。

手順2105では、実施の形態4の手順1311と同様に、顔決定手段2013が、左目位置信号2011と右目位置信号2012を読み込んで、顔の領域の座標を表す顔領域信号2014を出力する。

手順2106では、情報検出手段2015において、顔領域信号2014と画像信号2004を読み込んで、顔の領域の画像から顔の特徴量を求め、第1の顔特徴信号2016を出力する。顔の特徴量を求める方法としては、特願平11-295058に開示されている方法を利用することができる。この方法では、あらかじめ認識したい人達の顔を撮影したビデオカメラからの顔画像をセットA1（ビデオ顔画像）とし、認識したい人達の顔の映った身分証を撮影したイメージスキャナからの身分証の顔画像をセットB1（身分証写真）として、顔画像データを1次元のデータ列としたパターンAとパターンBに変換してから、学習によってセットA1用の特徴抽出行列とセットB1用の特徴抽出行列を計算しておく。そして、顔の領域の画像データに対してセットA1用の特徴抽出行列を適用することで、顔の特徴量を求める。

手順2107では、認識手段2018が、データベース手段2017に蓄積さ

れたすべての顔の特徴量を表す第2の顔特徴信号2019を読み取り、第1の顔特徴信号2016に類似した顔の特徴量を選び出す。そして、最も類似した顔の特徴量をもつ人物の名前をデータベース手段から読み取り、認識結果信号2021として出力する。データベース手段に蓄積しておく顔の特徴量は、顔画像セットB1に対してセットB1用の特徴抽出行列を適用して、あらかじめ求めておくものとする。

本実施の形態によれば、目の間は顔の表情による変動が少なく口やヒゲと離れているので、目や口の動きを含む顔の表情やヒゲの有無や髪型による顔の変動に強く、顔の領域を検出して高精度に認識することができる。

10 以上のように、本発明によれば、画像中から目の候補領域と目の間の候補領域を検出し、目の候補領域と目の間の候補領域の位置関係から目の間の領域を決定し、目の間の領域を含むように目や顔の領域を決定することで、目の間は顔の表情による変動が少なく、口やヒゲと離れているので、目や口の動きを含む顔の表情やヒゲの有無による顔の変動に強く、人物の顔の領域などの情報を検出することが
15 ができる。

また、複数の画素が連結領域かどうかを判定するのではなく、最初に小領域に分割してから小領域が目や目の間であるかどうかを判定することで、少ない計算量で実行することができる。

さらに、目の間の領域の重なりを調べたり、目の間の候補領域内の特徴や目の候補領域内の特徴を判定して目の間の領域を決定することで、顔以外の背景物が
20 多く映っている場合を含めて、少ない計算量で人物の顔の領域の誤検出を抑制し、高精度に検出することができる。

以上の説明では、コンピュータにプログラムを実行させることにより、入力画像の輝度特徴から目の候補領域と目の間の候補領域とを検出し、前記目の候補領域と目の間の候補領域との位置関係から、前記目の間の領域を検出し、検出され
25 た目の間の領域から顔の領域を決定する処理を実現している。このようなプログラムを記憶媒体に格納して販売、使用することができ、又は情報通信回線を介して配布することができる。また、入力画像の輝度特徴から目の間の領域を検出し、目の間の領域を含むように両目の候補領域を設定し、両目の候補領域の輝度特

徴または形状特徴を調べて右目の位置と左目の位置を決定し、右目の位置と左目の位置を含むように顔の領域を決定し、決定した顔の領域を出力するプログラムを記憶媒体に格納して販売したり、情報通信回線を介して配布するようにしてもよい。さらに、動画像のフレーム画像から上述した人物の顔の検出方法で検出し

- 5 た顔の領域から、検出された各領域の顔に関する情報を検出して蓄積し、前記顔に関する情報を用いて動画像を検索するプログラムを記憶媒体に格納して販売したり、情報通信回線を介して配布してもよい。さらに、上記した人物の顔の検出方法で検出した顔の領域から、検出された各領域の顔の特徴量と予めデータベースに蓄積された顔の特徴量とを用いて顔の認識を行うプログラムを記憶媒体に格納して販売したり、情報通信回線を介して配布するようにしてもよい。
- 10

本明細書は、2000年2月1日出願の特願2000-023680に基づくものである。その内容はここに含めておく。